BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

REPLICATING BY THE CLOCK

Alon Goren and Howard Cedar

The eukaryotic genome is divided into well-defined DNA regions that are programmed to replicate at different times during S phase. Active genes are generally associated with early replication, whereas inactive genes replicate late. This expression pattern might be facilitated by the differential restructuring of chromatin at the time of replication in early or late S phase.

G BANDS
A characteristic chromosome-banding pattern that is shown by staining with Giemsa. Light and dark G bands differ in their molecular and regulatory features, such as gene density, repetitive sequence elements and replication timing.

ISOCHORES
Long DNA fragments (>300 kb) defined by their average G+C content. Isochores are divided into five subfamilies according to their G+C composition: from G+C poor (~40% G+C) to G+C rich (~55–60% G+C).

Department of Cellular Biochemistry and Human Genetics, Hebrew University, Ein Kerem, Jerusalem 91120, Israel.
Correspondence to H.C.
e-mail: cedar@md.huji.ac.il
doi:10.1038/nrm1008

Animal-cell DNA is arranged in distinct chromosomal bands that undergo replication in a temporally programmed way during S phase[1]. Although the reason for this organization is not completely understood, there is no question that it has functional significance. A great deal of research has shown that, both at a regional and sequence-specific level, there is a striking and straightforward correlation between replication timing and gene expression. Most active genes replicate early (in the first half) in S phase, whereas many inactive gene sequences replicate late in S phase[2]. Despite this intriguing relationship, the mechanism that connects replication timing and gene expression has not yet been fully elucidated.

One of the most interesting suggestions for explaining this phenomenon is the 'window of opportunity' model[3,4]. According to this model, the transcriptional competence can be influenced at the time of replication, when the chromatin structure is disrupted and then repackaged on newly synthesized DNA. One possibility is that genes undergoing replication in early S phase are exposed to factors that are required for the formation of active transcription complexes, whereas genes that replicate in late S phase experience a different nuclear environment, which is more conducive for the generation of repressive structures. This model relies on three components: first, an independent mechanism for controlling replication timing; second, the susceptibility of chromatin to restructuring during replication; and third, variations in nuclear composition as a function of progression through S phase.

Replication time zones

Animal cells are characterized by a striking division of the genome into replication time zones. These are best visualized by labelling the cells for various time intervals with bromodeoxyuridine (BrdU) and examining their incorporation pattern in metaphase chromosomes[5] (FIG. 1). High-resolution analysis of prometaphase chromosomes indicates that the average length of each time zone is approximately 1 Mb[6,7]. Early-replicating regions map directly to light G BANDS, which represent the G+C-rich ISOCHORES[8], whereas the late-replicating regions map to dark G bands[1], which generally contain A+T-rich DNA[9,10]. So, replication timing seems to reflect a fundamental aspect of the genome sequence structure.

In contrast to bacteria, which replicate their entire genome using a single origin, animal cells copy their DNA by initiating bidirectional synthesis at multiple origins that are spaced, on average, ~100 kb from each other[11]. This indicates that each replication time band is made up of several replicons, which must fire in a coordinated manner — and it is exactly this pattern that is observed when DNA synthesis is visualized by autoradiography after labelling with H^3-thymidine[12]. A similar organization was observed using microarray technology to map the full complement of replication origins in the yeast *Saccharomyces cerevisiae*. In this simple organism, the genome seems to also be made up of small clusters that contain several origins which have coordinated replication-time properties[13]. The picture that emerges from these studies is that the control of replication timing is carried out at the regional level, and is executed by the coordinated firing of multiple origins.

The timing of replication has also been determined for individual genes, by isolating BrdU-labelled DNA from cells at different stages of S phase[14–18] and then

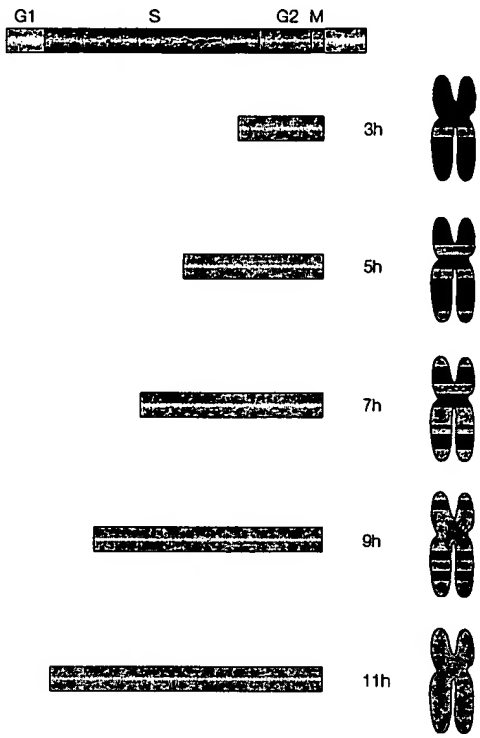


Figure 1 | **Replication bands.** In a typical cell cycle, S phase lasts about 8 h, whereas G2 takes about 2 h. Replication bands can be observed by labelling with bromodeoxyuridine (BrdU) (green) for various lengths of time and then examining only those cells that contain metaphase chromosomes. This functions as a form of retroactive synchronization. When labelled for 3 h, for example, the BrdU that is observed on chromosomes must have been incorporated into DNA during the last hour of S phase. Labelling for 5 h shows the DNA that was replicated during the last 3 h of S phase. The entire DNA is labelled when cells are incubated with BrdU for 11 h. BrdU labelling is regional and appears in a band-like pattern, which coincides with the replication time zones.

assaying for specific sequences. Early-replicating genes are concentrated in fractions from the beginning of S phase, whereas late-replicating sequences are only found in BrdU-labelled DNA from cells in the later stages of S phase.

Analysis of many genes has shown that there is a striking correlation between expression and early replication timing, but this relationship is certainly not absolute. Microarray analysis in *Drosophila melanogaster* has shown that, unlike most genes, there are clearly active gene sequences that are located in late-replicating regions, as well as inactive genes that replicate early in S phase[19]. Replication timing is also subject to developmental regulation. Housekeeping genes that are transcribed constitutively, for example, undergo replication relatively early in S phase, whereas many tissue-specific genes replicate late in most cell types, but become early replicating in the expressing tissue[2,20].

Unlike these higher organisms[21], there is no general correlation between replication timing and gene expression in *S. cerevisiae*[13], but some repressed genes are indeed located in subtelomeric regions that replicate late in S phase. The HML and HMR mating-type genes are a good example of this phenomenon. Although constitutively silent in their telomeric positions, these same sequences are actively expressed when they are copied into the early-replicating mating-type (MAT) locus on the same chromosome[22].

As DNA replication seems to be regulated at the regional level, it is interesting to address the question of how these topographical patterns are altered for developmentally regulated genes. This type of molecular-band analysis has been carried out using fluorescence *in situ* hybridization (FISH)[20], whereby specific sequences are visualized in interphase nuclei from non-synchronized cells by using fluorescent probes. Unreplicated loci appear as two single dots in diploid cells, whereas replicated DNA is characterized by double signals (FIG. 2a). A high percentage of double dots in S-phase nuclei indicates early replication, whereas numerous single signals is characteristic of late replication. This method is easy to carry out, and by using a series of probes that surround a specific gene, a good picture of the size of each replication time zone can be obtained.

Analysis of the cystic fibrosis gene (*CFTR*), for example, shows that this sequence is normally located in a large late-replicating region on human chromosome 7. In active cells, however, a domain of ~700 kb becomes early replicating, thereby generating a 'new' chromosomal mini-band[20]. A similar process takes place at the β-globin locus that is located on chromosome 11, but, in this case, it is a large region of >1 Mb that becomes early replicating in erythroid precursor cells[23]. So, it seems that the developmental control of replication timing is also carried out in a regional manner. What emerges from these studies is that chromosome bands represent a basic, but dynamic, feature of the genome that is characterized by specific structure[24], function and replication-timing properties.

Asynchronous replication timing

Although most sequences in the genome are expressed equally from both alleles, there is a small class of genes that are transcribed preferentially from a single allele in each cell. In the light of the strong correlation between gene expression and replication timing, it is interesting to ask whether these genes replicate in an allele-specific manner.

Genomic imprinting. Genomic imprinting is a classic example of asynchronous replication timing. In this case, it is always one parental allele, either the maternal or the paternal, that is expressed preferentially. One of the most important characteristics of imprinted genes is that they are always organized into clusters that occupy large chromosome domains. The regulation of imprinting and the coordination between individual genes in each domain are dictated at the regional level by means of imprinting control centres[25] (BOX 1).

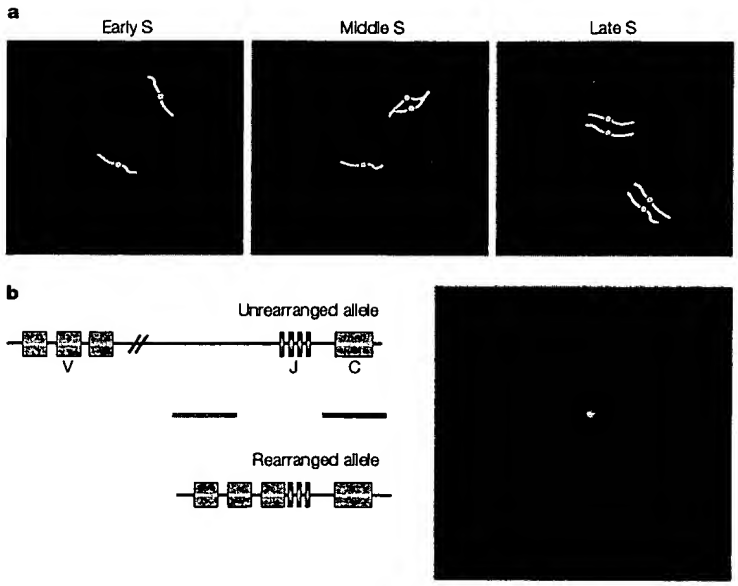


Figure 2 | **FISH analysis of monoallelically expressed genes. a** | Fluorescence *in situ* hybridization (FISH) is carried out by using fluorescently labelled probes for specific genes. Cells are labelled with bromodeoxyuridine (BrdU) (red) to analyse nuclei in S phase, exclusively. Two single dots (yellow) indicate that the gene has not yet replicated, whereas two double dots are seen in nuclei in which both alleles have already replicated. Hybridization dots and DNA molecules are shown schematically in this figure. Genes that replicate asynchronously show a high percentage (>30%) of nuclei with one single and one double dot. It should be noted that the FISH method can only detect replication after sister chromatid segregation, and therefore might not always represent an accurate indication of replication timing. **b** | The κ-locus replicates asynchronously in all cell types. In many B cells, one allele has undergone rearrangement with a single variable (V) region recombined to one junction (J) segment. The other allele is not rearranged. The nucleus shows diamidinophenylindole (DAPI) staining for DNA (blue) and FISH analysis with two separate probes, one from the constant (C) region (as indicated by the red line; red labelling) and the other from the region between V and J (as indicated by the yellow line; yellow labelling) that is eliminated during rearrangement. This nucleus shows asynchronous replication whereby only one allele has replicated and the other has not. The yellow signal identifies the unrearranged germline allele. In >90% of the cases, the single red dot is associated with the yellow FISH signal, which indicates that it is usually the early-replicating allele that undergoes rearrangement.

Although the exact mechanism of imprinting has not yet been deciphered, it is clear that each individual allele must aquire an EPIGENETIC mark while still in the gametes, and that this differential structure then has to be maintained in an autonomous manner throughout development. There is strong evidence that gamete-specific methylation has a crucial function in this process[26], but it is now clear that this modification alone cannot explain all aspects of imprinting, and that other epigenetic factors must be involved[27].

FISH analysis of interphase nuclei has been used to study replication timing at imprinted gene loci. For most regions in the genome, both alleles seem to replicate in a coordinated way, so that nuclei show mostly (>85%) either single/single or double/double signals for sets of alleles. By contrast, for imprinted loci, a large number of nuclei have a single dot on one allele and double dots on the other, which indicates that one allele has replicated before the other (FIG. 2a). This pattern of asynchronous replication is regional in nature and encompasses the entire imprinted domain[28,29].

Although there is no direct evidence that replication timing itself influences imprinted gene expression, there is good indirect developmental evidence that it is indeed important. By using FISH to examine replication timing during gametogenesis and early development, it was shown that parent-of-origin-specific replication timing is 'set up' around the time of meiosis, is maintained through development and is then erased early during the next round of gametogenesis[30]. This epigenetic feature therefore fulfils all of the criteria of a primary imprinting mark (BOX 1). Furthermore, as in almost all cases it is the paternal allele that replicates early, this also represents a mechanism by which the parental allele in each cell can be identified. Asynchronous replication timing is controlled by the same regulatory centres that define the imprinted chromatin and expression patterns of the two alleles[31,32], which indicates that it is indeed an integral part of the domain-wide imprinting process.

X-chromosome inactivation. Another form of monoallelic expression is exemplified by X-CHROMOSOME INACTIVATION in female mammals. In this process, one X chromosome in each cell is chosen randomly to undergo chromosome-wide inactivation in the late blastocyst, and this inactivation is then maintained in a clonal manner throughout development. As a result, every somatic cell has only one active X chromosome, either the maternal or the paternal. This process is directed *in cis* by the X-INACTIVATION CENTRE (Xic), which is responsible for a range of molecular events, including the synthesis of an X-inactivation specific transcript (Xist) and chromosome-wide epigenetic changes that involve chromatin structure, DNA methylation and replication timing[33]. Although the inter-relationship between these epigenetic marks is not well understood, it is worth noting that the shift to late replication timing is one of the first developmental changes that is associated with the inactivation process, which occurs as early as several days before *de novo* methylation in the mouse[34].

Allelic exclusion. In addition to sequences on the X chromosome, several other genes have also been found to have a random monoallelic expression pattern. In these cases, which include immune-system- and olfactory-receptor genes, the regulation process is also regional and characterized by asynchronous replication[35,36], and, in this sense, they behave in a similar way to the X chromosome. Both alleles replicate in an apparently synchronous manner at early stages of embryogenesis, but at the time of implantation, one allele becomes early replicating and the other late replicating in each individual cell. This pattern is maintained in a clonal manner throughout development, with the paternal allele replicating early in some cells and late in others[30,37].

The immune-system-receptor genes are a good example of how this epigenetic mark can influence function. During lymphoid development, designated B- and T-cell-receptor loci undergo rearrangement as a

EPIGENETIC
Any heritable influence on the function of a chromosome or gene that is not caused by a change in DNA sequence.

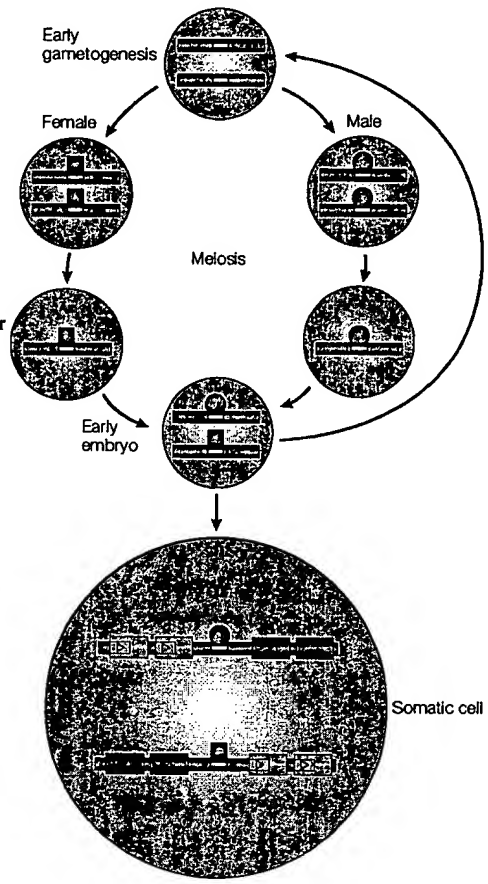
X-CHROMOSOME INACTIVATION
The transcriptional inactivation of one of the two X chromosomes in female embryos. The choice of the maternal or the paternal allele is random, but is then maintained clonally in subsequent cell generations. The inactive X chromosome is characterized by DNA methylation, late replication timing and condensed chromatin structure.

Box 1 | Genomic imprinting

Genomic imprinting was originally discovered from nuclear transplantation experiments in mice, which showed that development can only proceed normally if the embryo is constructed from both maternal and paternal genome components[77,78]. This effect evidently comes about because some genes are only expressed from a single allele, either maternal or paternal. Further studies showed that these genes are organized into clusters that are subject to coordinated control[25].

Genomic imprinting is controlled by parent-of-origin-specific epigenetic marks. Imprinted domains contain an imprinting centre (white), which is marked during late gametogenesis to identify it as either paternal (purple circle) or maternal (red square). Markings might include DNA methylation, replication timing and chromatin structure. After fertilization, these epigenetic states are stable and are maintained through cell division, but they are erased in early gametic cells that are derived from the embryo. After implantation, the imprinting centre can induce additional epigenetic signals on the domain, and this brings about an imprinted expression pattern, with some genes being specifically expressed (yellow) on the maternal allele (pink) and others on the paternal allele (green).

Imprinting is also of medical significance, and several genetic diseases are associated with defects in imprinted gene regions. Chromosome 15, for example, contains a large cluster of paternally expressed imprinted genes[79], and deletion of the paternal allele causes Prader–Willi syndrome, a developmental disease that is characterized by defects in growth control and brain function. Deletion of the maternal allele, on the other hand, brings about Angelman syndrome, which is a completely different disease that is associated with severe mental retardation. In rare cases, these diseases can actually come about through the disruption of the imprinting centre itself, which leads to misregulation of the entire imprinted domain[79].



mechanism for generating the B- and T-cell-receptor repertoires. They do this by choosing individual variable (V) and diversity (D) genes from a multigenic cluster, thereby assembling a specific immune-receptor molecule[38] (FIG. 2b). Demethylation and chromatin opening are required to make these regions accessible to the rearrangement machinery, and occur initially on only a single allele in each cell[39]. This might be one of the key mechanisms for achieving ALLELIC EXCLUSION[40]. Strikingly, in almost all instances, this first rearrangement event takes place on the early-replicating allele (FIG. 2b), which indicates that this decision is actually programmed into the immune-receptor loci early in development.

Olfaction is another system that involves the selection, in each cell, of a single candidate from a multigenic array. More than 1,000 individual olfactory-receptor genes are organized into numerous clusters that are scattered throughout the genome. In each olfactory neuron, however, only a single gene is actually expressed on its surface[41]. This is carried out through a complex process that stochastically selects a single gene from one developmentally predetermined cluster, and only one parental allele, either the maternal or the paternal, is expressed in each cell[35]. Although the mechanism for this type of allelic exclusion is not known, it should be noted that all olfactory-receptor loci have been shown to undergo random asynchronous replication, and this might reflect inherent differences between the two alleles in each cell. The choice between X-chromosome-linked green or red pigment genes in cone cells of the retina might also operate by a similar mechanism[42]. So, it seems that asynchronous replication timing reflects a basic difference in chromatin structure between the two alleles. This epigenetic marker is set up early in development, maintained in a clonal manner and can then be used as a mechanism of allelic exclusion during terminal differentiation.

As many regions of the genome undergo random asynchronous replication, it is interesting to ask whether this process is coordinated, and if so, how? To this end, several gene regions were analysed for replication timing in monoclonal cell lines. Four individual markers that are scattered over the full length of chromosome 6 were all found to replicate in a coordinated manner; some clones had all four markers replicating

X-INACTIVATION CENTRE (Xic). A *cis*-acting region on the X chromosome that produces the X-inactivation-specific transcript (Xist) that is necessary for initiating X-chromosome inactivation in female cells.

ALLELIC EXCLUSION The process by which a cell (for example, from the immune or olfactory system) uses either the gene from its maternal chromosome or the one from the paternal chromosome, but not both.

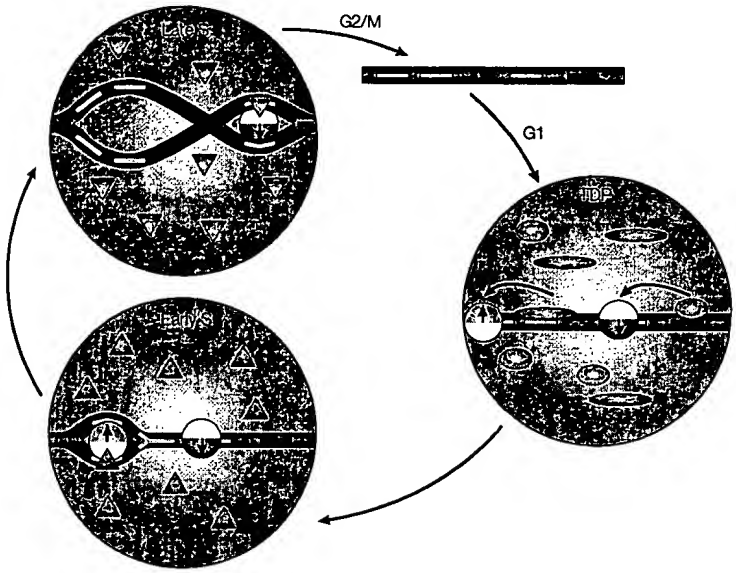


Figure 3 | **Control of replication timing.** A DNA region with two origins (white) is shown. One has *cis*-acting sequences that direct early replication (yellow) and the other has a *cis*-acting element that directs late replication (red). At the time of mitosis, the nuclear membrane is disrupted, and it then takes some time before the nucleus becomes reorganized in early G1. At the timing decision point (TDP), *trans*-acting factors (pink and green ellipses) interact with the *cis*-acting sequences to induce either an early (yellow 'clock') or a late (red 'clock') epigenetic state on the origins. This epigenetic mark could be in the form of nuclear localization or histone modification. Once the origin is marked, these *cis*-acting sequences and *trans*-acting factors are no longer required. During early S phase, the early origin is identified by early-S-phase-specific *trans*-acting factors (blue triangles) that cause it to replicate. In late S phase, a different set of *trans*-acting factors (orange triangles) recognizes the late origins and causes them to fire. The epigenetic states of early and late origins are then erased before the next G1 stage of the cell cycle.

early on the maternal allele, whereas others had all four markers replicating early on the paternal allele. Asynchronous replicating regions on other chromosomes also showed internal coordination, but different chromosomes behaved independently[80]. These findings indicate that each autosome might have one or more central control elements that function to coordinate several asynchronously replicating domains *in cis*, in a manner that is similar to the Xic.

Cis-acting regulatory sequences

Although the replication time of any genomic sequence seems to be a function of both the firing time and its distance from the nearest origin — that is, late replication can occur from early-firing origins[43] — these elements themselves do not usually have any function in actually setting the replication clock. In yeast, almost all of the authentic origins of replication have been identified[13], and some have been subjected to genetic analysis. By moving a late-replicating origin away from its subtelomeric position, it is able to initiate replication in early S phase[44]. Conversely, the placement of an early-replicating origin near the telomere automatically dictates that it will fire late in S phase, in both yeast[45] and human cells[46].

These experiments indicate clearly that replication timing is directed by *cis*-acting elements that can exert control over origin sequences that are located in their domain of influence, in a way that is possibly similar to how enhancers activate gene promoters. As many of the late-replicating regions in the yeast genome are actually subtelomeric, it is probable that sequences in the telomere itself are important. Indeed, studies with episomal plasmids have shown that even short $C_{1-3}A$ tracts can drive late replication from a test origin[45]. Furthermore, genetic experiments indicate that this might be mediated by the telomere-binding proteins Ku[47] and Sir3 (REF. 48). In contrast to this relatively simple system, late replication of a multi-origin, 130-kb non-telomeric region on *S. cerevisiae* chromosome XIV is directed by a combination of at least three sequence elements that seem to work in concert[49].

The replication origins in animal cells are also subject to developmental control. The human and mouse β-globin loci, for example, undergo bidirectional replication from origin sequences that are located upstream of the β-globin gene[50–54]. The same origin region is used in both non-erythroid cells, which replicate this region late in S phase, and in erythroid cells, in which the locus undergoes replication in early S phase. Experiments in human cells from patients with hispanic thallassaemia indicated initially that this process is regulated by *cis*-acting sequences that are located within 40 kb surrounding the LOCUS CONTROL REGION (LCR)[55]. It has now been shown definitively that replication timing in this region is controlled by several individual elements both within[23] and outside[56] the LCR.

It seems that neither early nor late replication is the default state at this locus. Rather, the LCR operates more like a toggle switch that can direct, in a dominant manner, either early replication in erythroid cells or late replication in other tissues[23]. This is undoubtedly mediated by tissue-specific *trans*-acting factors, because the introduction of a normally late-replicating β-globin locus into an erythroid environment is sufficient to bring about early replication of the entire region[55]. The fact that the LCR is a multifunctional effector that is responsible for the developmental regulation of gene expression, chromatin structure and replication timing indicates clearly that all of these parameters are intimately related and centrally controlled.

Control of replication timing

The replication-timing process seems to involve three key components: the epigenetic marking of origins as either early or late; the use of cell-cycle-regulated protein factors to activate origins at the proper times in S phase; and a system of CHECKPOINT genes that ensure the replication timing programme is carried out according to schedule.

Epigenetic marking of replication origins. Although replication itself takes place exclusively during S phase, the programming of this process begins in the G1 stage of the cell cycle when each origin is made competent for firing

LOCUS CONTROL REGION (LCR). A large regulatory sequence that harbours several elements that control gene expression and chromatin structure during development.

CHECKPOINT
A point at which the cell-division cycle can be halted until conditions are suitable for the cell to proceed to the next stage.

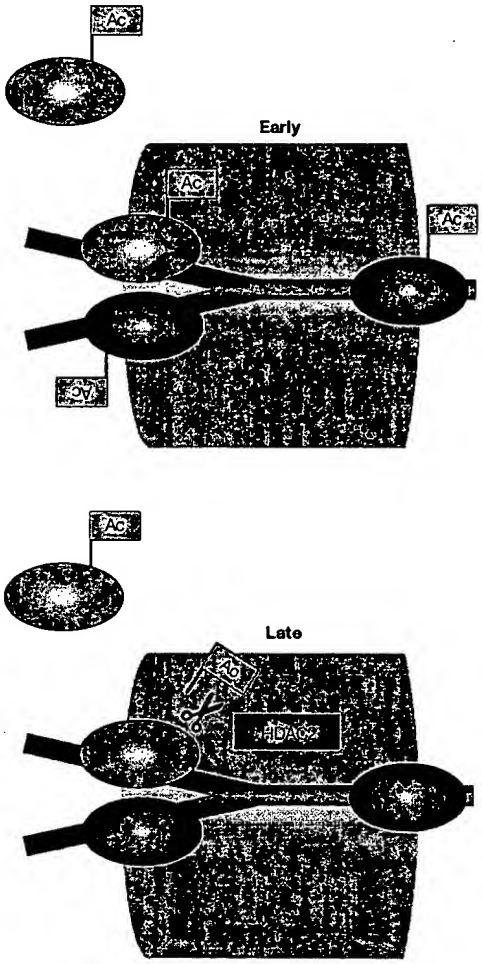


Figure 4 | **Chromatin formation in early and late S phase — a model.** During replication, new nucleosomes (pink) are assembled on one molecule of newly synthesized DNA, whereas the other retains old nucleosomes (brown). The new nucleosomes might already be pre-acetylated. In early S phase, the replication machinery (blue) assembles DNA using pre-made nucleosomes (already acetylated). In late S phase, the histone deacetylase HDAC2 is present at replication foci, and this causes local deacetylation[72]. For this reason, late-replicating DNA is assembled with underacetylated histones. It should be noted that this model is based on experiments that were carried out in tissue culture, and further studies will be needed to confirm this mechanism and establish its actual contribution to gene repression *in vivo*.

through the formation of a pre-replication complex. The fact that this can only be accomplished in G1 (REF. 57) is an important aspect of replication control, because it ensures that each origin can only fire once per cycle. In a similar way, the instructions for establishing the time when each origin will fire in S phase are also provided during G1. This was shown in a set of experiments in yeast, in which recombinase recognition sequences were introduced to flank a late-replicating origin that was located adjacent to the telomere. By inducing recombinase activity at specific times, the origin fragment could be released from the telomeric regulatory sequences at different pre-determined points in the cell cycle. The occurrence of recombination after the cell had passed through the first few hours of G1 produced an episomal DNA fragment that was already committed to replicate late in S phase. By contrast, detachment before G1 caused the loss of control and the resulting plasmid reverted to early replication[44]. These results indicate that *cis*-acting sequences are needed in early G1 to set up the replication timing profile, but, once this has been done, these elements are no longer required (FIG. 3).

Animal cells might use a similar strategy for controlling replication timing. This was shown by assaying the replication properties of pre-labelled animal-cell nuclei that were placed in *Xenopus* egg extracts. Nuclei taken from cells in late G1 phase undergo early and late replication in accordance with their pre-programmed schedule. By contrast, nuclei that have not yet passed through a crucial timing decision point (TDP) at the beginning of G1 are unable to direct organized replication timing[57].

Differences in replication timing seem to be associated with major variations in nuclear organization. Early replication always takes place at distinctive loci that are dispersed throughout the interphase nucleus. By contrast, replication during middle S phase is associated with the nuclear periphery, whereas late-replication foci colocalize with HETEROCHROMATIN[58]. It is striking that these structural features seem to be set up together with the replication time profile in early G1 (REFS 57, 59). Although this is suggestive, there is as yet no direct evidence that the establishment of replication timing is mediated by the nuclear restructuring that occurs after each passage of the cell through mitosis. Indeed, it has been shown both in yeast and in animal cells that, once set up, the replication timing instructions remain fixed, even if there is a change in nuclear positioning[60,61].

Local chromatin structure might also be involved in the determination of replication timing. In yeast, for example, early replicating origins seem to be packaged together with acetylated histones, whereas late-replicating origins are part of a deacetylated structure. Furthermore, when a HISTONE ACETYLTRANSFERASE (HAT) was recruited artificially to a late-replicating origin, this was sufficient to shift the replication event to an earlier time in S phase[62]. In animal cells, treatment with a histone deacetylase inhibitor also affects the replication timing at specific loci[63]. These studies indicate clearly that histone acetylation is one of the epigenetic marks that direct replication timing, although others, including DNA methylation, might also be important in this process[64,65].

Activation and regulation of origin firing. In the next step of replication-timing control, the signals that mark replication origins as being early or late must be read and interpreted during S phase. Judging by other cell-cycle-controlled processes, this might require the generation of a gradient over time, and of protein factors that can recognize the predetermined epigenetic pattern that is set up during G1 (FIG. 3). It has been

HETEROCHROMATIN
A cytologically defined genomic component that contains repetitive DNA (highly repetitive satellite DNA, transposable elements and ribosomal DNA gene clusters) and some protein-encoding genes.

HISTONE ACETYLTRANSFERASE (HAT). An enzyme that modifies histone tails covalently by adding an acetyl group to lysine residues, thereby changing their structure.

CHROMATIN IMMUNOPRECIPITATION (ChIP). A technique that isolates sequences from soluble DNA chromatin extracts (complexes of DNA and protein) by using antibodies that recognize specific chromosomal proteins.

CPG ISLANDS Sequences (0.5–2 kb) that are rich in the CpG dinucleotide, which are mostly located upstream to housekeeping and some tissue-specific genes. They are constitutively non-methylated in all animal cell types.

shown that cyclins themselves are implicated in the control of replication timing. So, in yeast, for example, the B cyclin Clb5 is required specifically for the replication of late origins, and, in its absence, the entire genome is copied using early-firing origins exclusively [66]. Rad53 and Mec1 seem to be used to coordinate the normal process of replication-timing control by delaying late-origin firing until the early replicons have completed their process of elongation[67,68]. As such, this is an important component of the intra-S-phase checkpoint mechanism[69].

Effect of replication timing on gene expression

There is now extensive evidence that indicates that the genome-wide pattern of replication timing is regulated in an independent and programmed manner, and is not merely a secondary consequence of gene expression[23,56]. This raises the key question of whether replication timing itself can affect transcription. As proposed by the 'window of opportunity' model, replication causes the disruption of chromatin structure and, as a result, provides a unique opportunity for nuclear factors to interact directly with DNA sequences, thereby influencing the transcriptional competency[70]. As the protein environment in the nucleus most certainly changes with progression through S phase, replication timing could have a crucial function in establishing the expression profile of the genome.

A basic prediction of this model would be that reporter genes should undergo transcription more efficiently when introduced into an early-S-phase, as opposed to a late-S-phase, nucleus. This idea has recently been tested by single-cell microinjection experiments[71]. Reporter genes that were injected into early-S-phase nuclei were almost tenfold more transcriptionally active than if the same gene was injected into late-S-phase nuclei. And, once established, these transcriptional states remain stable when the cell continues cycling. As similar results were obtained with various promoter sequences, this effect probably cannot be attributed to the involvement of transcription factors. Instead, it now seems that this S-phase specificity is probably mediated by chromatin structure. Indeed, CHROMATIN IMMUNOPRECIPITATION (ChIP) analysis shows that early-injected DNA is packaged into chromatin that contains acetylated histones, whereas late-injected templates are underacetylated[71].

Evidence exists that shows a correlation between early replication and open chromatin structure, as determined by DNase I sensitivity[24] and histone acetylation[72]. The injection experiments described above now indicate that replication timing itself might have a role in establishing this genome-wide pattern. The mechanism for this process could involve the histone deacetylase HDAC2, which has been shown to localize exclusively to late-S-phase replication foci[73]. So, although all newly replicated DNA might be packaged initially with acetylated histone H4 (REFS 74,75), nucleosomes that are assembled in late S phase could be deacetylated specifically. Given that the replication-timing programme is reset automatically at the beginning of each cell cycle, this might represent a simple system for maintaining chromatin states in dividing cells (FIG. 4).

Role of replication-timing control

Unlike the genome of simple organisms, in which almost all of the genes are expressed ubiquitously, the animal genome contains numerous gene sequences that are repressed in most cell types. This process is controlled in a complex way by a combination of sequence-specific and global repression mechanisms. In the case of DNA methylation, for example, all non-CPG ISLAND sequences are methylated automatically early in development regardless of their sequence. This pattern brings about a closed chromatin conformation, which can be passed on from generation to generation by means of a semi-conservative maintenance mechanism[76]. So, it seems that DNA replication timing might work in a similar way. Regions that contain constitutively active genes are programmed to be early replicating, whereas the rest of the genome replicates late, thereby creating a global pattern of repression that can be regenerated during each cell cycle. Furthermore, replication timing can also be regulated, and, in this way, allows tissue-specific genes to undergo programmed activation during development.

Perspectives

There are three potential ways to maintain gene repression in dividing cells. First, the most direct mechanism is genetic in nature and involves DNA sequence elements that function to recruit *trans*-acting factors to specific regions of the genome after each round of replication. Second, DNA methylation, which inhibits gene expression by altering the chromatin structure, is a semi-genetic mechanism; although it is not based directly on sequence information, it does involve covalent changes in DNA structure. Third, replication-timing control seems to constitute another mechanism for maintaining functional states by directing the restructuring of chromatin after DNA replication. This pathway is truly epigenetic, as, once established, it is completely independent of local DNA sequences.

1. Hand, R. Eucaryotic DNA: organization of the genome for replication. *Cell* **15**, 317–325 (1978).
2. Holmquist, G. P. Role of replication time in the control of tissue specific gene expression. *Am. J. Hum. Genet.* **40**, 151–173 (1987).
3. Gottesfeld, J. & Bloomer, L. S. Assembly of transcriptionally active 5S RNA gene chromatin *in vitro*. *Cell* **28**, 781–791 (1982).
4. Brown, D. D. The role of stable complexes that repress and activate eucaryotic genes. *Cell* **37**, 359–365 (1984).
5. Latt, S. A. Microfluorometric detection of deoxyribonucleic acid replication in human metaphase chromosomes. *Proc. Natl Acad. Sci. USA* **70**, 3395–3399 (1973).
6. Drouin, R., Lemieux, N. & Richer, C. L. Analysis of DNA replication during S phase by means of dynamic chromosome banding at high resolution. *Chromosoma* **99**, 272–280 (1990).
7. Yunis, J. J. Mid-prophase human chromosomes. The attainment of 2000 bands. *Hum. Genet.* **56**, 293–298 (1981).
8. Bernardi, G. Isochores and the evolutionary genomics of vertebrates. *Gene* **241**, 3–17 (2000).
9. Tenzen, T. *et al.* Precise switching of DNA replication timing in the GC content transition area in the human major histocompatibility complex. *Mol. Cell. Biol.* **17**, 4043–4050 (1997).
10. Watanabe, Y. *et al.* Chromosome-wide assessment of replication timing for human chromosomes 11q and 21q: disease-related genes in timing-switch regions. *Hum. Mol. Genet.* **11**, 13–21 (2002).

11. Huberman, J. A. & Riggs, A. D. On the mechanism of DNA replication in mammalian chromosomes. *J. Mol. Biol.* **32**, 327–341 (1968).
12. Hand, R. Regulation of DNA replication on subchromosomal units of mammalian cells. *J. Cell Biol.* **64**, 89–97 (1975).
13. Raghuraman, M. K. *et al.* Replication dynamics of the yeast genome. *Science* **294**, 115–121 (2001).
The first report of a genome-wide replication timing map for a eukaryotic organism, the yeast *Saccharomyces cerevisiae*, which shows that there is no general correlation between gene expression and replication timing.
14. Goldman, M. A., Holmquist, G. P., Caston, L. A. & Nag, A. Replication timing of genes and middle repetitive sequences. *Science* **224**, 686–692 (1984).
15. Schmidt, M. & Migeon, B. R. Asynchronous replication of homologous loci on human active and inactive X chromosome. *Proc. Natl Acad. Sci. USA* **87**, 3685–3689 (1990).
16. Braunstein, D., Schulze, D., DelGiudice, T., Furst, A. & Schildkraut, C. L. The temporal order of replication of murine immunoglobulin heavy chain constant region sequences corresponds to their linear order in the genome. *Nucleic Acids Res.* **10**, 6887–6902 (1982).
17. Gilbert, D. M. Temporal order of replication of *Xenopus laevis* 5S ribosomal RNA genes in somatic cells. *Proc. Natl Acad. Sci. USA* **83**, 2924–2928 (1986).
18. Hansen, R. S., Canfield, T. K., Lamb, M. M., Gartler, S. M. & Laird, C. D. Association of fragile X syndrome with delayed replication of the *FMR1* gene. *Cell* **73**, 1403–1409 (1993).
19. Schubeler, D. *et al.* Genome-wide DNA replication profile for *Drosophila melanogaster*: a link between transcription and replication timing. *Nature Genet.* **32**, 438–442 (2002).
Genome-wide mapping that shows a strong correlation between replication timing and gene expression in *Drosophila melanogaster*.
20. Selig, S., Okumura, K., Ward, D. C. & Cedar, H. Delineation of DNA replication time zones by fluorescence *in situ* hybridization. *EMBO J.* **11**, 1217–1225 (1992).
21. Gilbert, D. M. Replication timing and metazoan evolution. *Nature Genet.* **32**, 336–337 (2002).
22. Reynolds, A. E., McCarroll, R. M., Newlon, C. S. & Fangman, W. L. Time of replication of ARS elements along yeast chromosome III. *Mol. Cell. Biol.* **9**, 4488–4494 (1989).
23. Simon, I. *et al.* Developmental regulation of DNA replication timing at the human β globin locus. *EMBO J.* **20**, 6150–6157 (2001).
24. Kerem, B. S., Goitein, R., Diamond, G., Cedar, H. & Marcus, M. Mapping of DNase-I sensitive regions on mitotic chromsomes. *Cell* **38**, 493–499 (1984).
25. Reik, W. & Walter, J. Genomic imprinting: parental influence on the genome. *Nature Rev. Genet.* **2**, 21–32 (2001).
26. Ferguson-Smith, A. C. & Surani, M. A. Imprinting and the epigenetic asymmetry between parental genomes. *Science* **293**, 1086–1089 (2001).
27. Rand, E. & Cedar, H. Regulation of imprinting: a multi-tiered process. *J. Cell. Biol.* (in the press).
28. Kitsberg, D., Selig, S., Keshet, I. & Cedar, H. Replication structure of the human β-globin gene domain. *Nature* **366**, 588–590 (1993).
29. Knoll, J. H. M., Cheng, S.-D. & Lalande, M. Allele specificity of DNA replication timing in the Angelman/Prader–Willi syndrome imprinted chromosomal region. *Nature Genet.* **6**, 41–46 (1994).
30. Simon, I. *et al.* Asynchronous replication of imprinted genes is established in the gametes and maintained during development. *Nature* **401**, 929–932 (1999).
Shows evidence that asynchronous replication timing in mammalian cells has the characteristics of a primary imprinting mark.
31. Shemer, R. *et al.* The imprinting of the Prader–Willi/Angelman syndrome domain. *Nature Genet.* **26**, 440–443 (2000).
32. Gunaratne, P. H., Nakao, M., Ledbetter, D. H., Sutcliffe, J. S. & Chinault, A. C. Tissue-specific and allele-specific replication timing control in the imprinted human Prader–Willi syndrome region. *Genes Dev.* **9**, 808–820 (1995).
33. Avner, P. & Heard, E. X-chromosome inactivation: counting, choice and initiation. *Nature Rev. Genet.* **2**, 59–67 (2001).
34. Lock, L. F., Takagi, N. & Martin, G. R. Methylation of the *HPRT* gene on the inactive X occurs after chromosome inactivation. *Cell* **48**, 39–46 (1987).
35. Chess, A., Simon, I., Cedar, H. & Axel, R. Allelic inactivation regulates olfactory receptor gene expression. *Cell* **78**, 823–834 (1994).
36. Bix, M. & Locksley, R. M. Independent and epigenetic regulation of the interleukin-4 alleles in CD4$^+$ T Cells. *Science* **281**, 1352–1354 (1998).
37. Mostoslavsky, R. *et al.* Asynchronous replication and allelic exclusion in the immune system. *Nature* **414**, 221–225 (2001).
38. Rajewsky, K. Clonal selection and learning in the antibody system. *Nature* **381**, 751–758 (1996).
39. Goldmit, M., Schlissel, M., Cedar, H. & Bergman, Y. Differential accessibility at the κ chain locus plays a role in allelic exclusion. *EMBO J.* **21**, 5255–5261 (2002).
40. Mostoslavsky, R. *et al.* κ chain monoallelic demethylation and the establishment of allelic exclusion. *Genes Dev.* **12**, 1801–1811 (1998).
41. Mombaerts, P. How smell develops. *Nature Neurosci.* **4**, 1192–1198 (2001).
42. Nathans, J. *et al.* Molecular genetics of human blue cone monochromacy. *Science* **245**, 831–838 (1989).
43. Ermakova, O. V. *et al.* Evidence that a single replication fork proceeds from early to late replicating domains in the IgH locus in a non-B cell line. *Mol. Cell* **3**, 321–330 (1999).
44. Raghuraman, M. K., Brewer, B. J. & Fangman, W. L. Cell cycle-dependent establishment of a late replication program. *Science* **276**, 806–809 (1997).
Shows that the replication timing profile is established by *cis*-acting elements during early G1 phase.
45. Ferguson, B. M. & Fangman, W. L. A position effect on the time of replication origin activation in yeast. *Cell* **68**, 333–339 (1992).
The first evidence that *cis*-acting sequences have a function in dictating the time of origin firing.
46. Ofir, R., Wong, A. C., McDermid, H. E., Skorecki, K. L. & Selig, S. Position effect of human telomeric repeats on replication timing. *Proc. Natl Acad. Sci. USA* **98**, 11434–11439 (1999).
47. Cosgrove, A. J. *et al.* Ku complex controls the replication time of DNA in telomere regions. *Genes Dev.* **16**, 2485–2490 (2002).
48. Stevenson, J. B. & Gottschling, D. E. Telomeric chromatin modulates replication timing near chromosome ends. *Genes Dev.* **13**, 146–151 (1999).
49. Friedman, K. L. *et al.* Multiple determinanats controlling activation of yeast replication origins late in S phase. *Genes Dev.* **10**, 1595–1607 (1996).
50. Kitsberg, D. *et al.* Allele-specific replication timing of imprinted gene regions. *Nature* **364**, 459–463 (1993).
51. Aladjem, M. I. *et al.* Participation of human β-globin locus control region in initation of DNA replication. *Science* **270**, 815–819 (1995).
52. Aladjem, M. I., Rodewald, L. W., Kolman, J. L. & Wahl, G. M. Genetic dissection of a mammalian replicator in the human β-globin locus. *Science* **281**, 1005–1009 (1998).
53. Aladjem, M. I. *et al.* Replication initiation patterns in the β-globin loci of totipotent and differentiated murine cells: evidence for multiple initiation regions. *Mol. Cell. Biol.* **22**, 442–452 (2002).
54. Kamath, S. & Leffak, M. Multiple sites of replication initiation in the human β-globin gene locus. *Nucleic Acids Res.* **29**, 809–817 (2001).
55. Forrester, W. C. *et al.* A deletion of the human β-globin locus activation region causes a major alteration in chromatin structure and replication across the entire β-globin locus. *Genes Dev.* **4**, 1637–1649 (1990).
56. Cimbora, D. M. *et al.* Long-distance control of origin choice and replication timing in the human β-globin locus are independent of the locus control region. *Mol. Cell. Biol.* **20**, 5581–5591 (2000).
57. Dimitrova, D. S. & Gilbert, D. M. The spatial position and replication timing of chromosomal domains are both established in early G1 phase. *Mol. Cell* **4**, 983–993 (1999).
The incubation in *Xenopus* egg extracts of mammalian nuclei isolated at different phases of the cell cycle showed that the organization of replication timing is established in early G1.
58. Gilbert, D. M. Nuclear position leaves its mark on replication timing. *J. Cell Biol.* **152**, F11–F15 (2001).
59. Li, F. *et al.* The replication timing program of the Chinese hamster β-globin locus is established coincident with its repositioning near peripheral heterochromatin in early G1 phase. *J. Cell Biol.* **154**, 283–292 (2001).
60. Bridger, J. M., Boyle, S., Kill, I. R. & Bickmore, W. A. Re-modelling of nuclear architecture in quiescent and senescent human fibroblasts. *Curr. Biol.* **10**, 149–152 (2000).
61. Heun, P., Laroche, T., Raghuraman, M. K. & Gasser, S. M. The positioning and dynamics of origins of replication in the budding yeast nucleus. *J. Cell Biol.* **152**, 385–400 (2001).
62. Vogelauer, M., Rubbi, L., Lucas, I., Brewere, B. J. & Grunstein, M. Histone acetylation regulates the time of replication origin firing. *Mol. Cell* **10**, 1223–1233 (2002).
Shows that in yeast the histone acetylation status in the vicinity of origins can determine its replication time.
63. Kagotani, K. *et al.* Replication timing properties within the mouse distal chromosome 7 imprinting cluster. *Biosci. Biotechnol. Biochem.* **66**, 1046–1051 (2002).
64. Hansen, R. S. *et al.* Escape from gene silencing in ICF syndrome: evidence for advanced replication time as a major determinant. *Hum. Mol. Genet.* **9**, 2575–2587 (2000).
65. Jablonka, E., Goitein, R., Marcus, M. & Cedar, H. DNA hypomethylation causes an increase in DNase-I sensitivity and an advance in the time of replication of the entire inactive X chromosome. *Chromosoma* **93**, 152–156 (1985).
66. Donaldson, A. D. *et al.* CLB5-dependent activation of late replication origins in *S. cerevisiae*. *Mol. Cell* **2**, 173–182 (1998).
67. Shirahige, K. *et al.* Regulation of DNA-replication origins during cell-cycle progression. *Nature* **395**, 618–621 (1998).
68. Santocanale, C. & Diffley, J. F. A Mec1- and Rad53-dependent checkpoint controls late-firing origins of DNA replication. *Nature* **395**, 615–618 (1998).
69. Paulovich, A. G. & Hartwell, L. H. A checkpoint regulates the rate of progression through S phase in *S. cerevisiae* in response to DNA damage. *Cell* **82**, 841–847 (1995).
70. Gilbert, D. M. Replication timing and transcriptional control: beyond cause and effect. *Curr. Opin. Cell Biol.* **14**, 377–383 (2002).
71. Zhang, J., Feng, X., Hashimshony, T., Keshet, I. & Cedar, H. The establishment of transcriptional competence in early and late S-phase. *Nature* **420**, 198–202 (2002).
Nuclear injection experiments showing that replicating DNA is automatically packaged in a repressed form during late S phase.
72. Jeppesen, P. & Turner, B. M. The inactive X chromosome in female mammals is distinguished by a lack of histone H4 acetylation, a cytogenetic marker for gene expression. *Cell* **74**, 281–289 (1993).
73. Rountree, M. R., Bachman, K. E. & Baylin, S. B. DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. *Nature Genet.* **25**, 269–277 (2000).
This work shows that histone deacetylase HDAC2 localizes to replication foci exclusively during late S phase in mammalian cells. This could function as a mechanism for repressing late-replicating genes.
74. Grunstein, M. Histone acetylation in chromatin structure and transcription. *Nature* **389**, 349–352 (1997).
75. Sobel, R. E., Cook, R. G., Perry, C. A., Annunziato, A. T. & Allis, C. D. Conservation of deposition-related acetylation sites in newly synthesized histones H3 and H4. *Proc. Natl Acad. Sci. USA* **92**, 1237–1247 (1995).
76. Razin, A. CpG methylation, chromatin structure and gene silencing — a three-way connection. *EMBO J.* **17**, 4905–4908 (1998).
77. Surani, M. A., Barton, S. C. & Norris, M. L. Development of mouse eggs suggests imprinting of the genome during gametogenesis. *Nature* **308**, 548–550 (1984).
78. McGrath, J. & Solter, D. Complementation of mouse embryogenesis requires both maternal and paternal genomes. *Cell* **37**, 179–183 (1984).
79. Schumacher, A. Mechanisms and brain specific consequences of genomic imprinting in Prader–Willi and Angelman syndrome. *Gene Funct. Dis.* **1**, 1–19 (2001).
80. Singh, N. *et al.* Coordination of the random asynchronous replication of autosomal loci. *Nature Genet.* (in the press).

Acknowledgements

This work was supported by grants from the National Institutes of Health, the Israel Cancer Research Fund and the Israel Science Foundation.

Online links

DATABASES

The following terms in this article are linked online to:

LocusLink: http://www.ncbi.nlm.nih.gov/LocusLink/
β-globin | *CFTR*
OMIM: http://www.ncbi.nlm.nih.gov/Omim/
Angelman syndrome | Prader–Willi syndrome
***Saccharomyces* Genome Database:**
http://genome-www.stanford.edu/Saccharomyces/
HML | HMR | MAT
Swiss-Prot: http://www.expasy.ch/
Clb5 | HAT | HDAC2 | Ku | Mec1 | Rad53 | Sir3

FURTHER INFORMATION

Howard Cedar's laboratory:
http://info.md.huji.ac.il/depts/humangenetics/cedar/cedar.html
Access to this interactive links box is free online.

Coordination of the random asynchronous replication of autosomal loci

Published online 10 February 2003; doi:10.1038/ng1102

Random monoallelic expression and asynchronous replication define an unusual class of autosomal mammalian genes. We show that every cell has randomly chosen either the maternal or paternal copy of each given autosome pair, such that alleles of these genes scattered across the chosen chromosome replicate earlier than the alleles on the homologous chromosome. Thus, chromosome-pair non-equivalence, rather than being limited to X-chromosome inactivation, is a fundamental property of mouse chromosomes.

Monoallelically expressed genes fall into three distinct classes. X inactivation in female cells is a random process resulting in half of the cells choosing the maternal X chromosome and half choosing the paternal X chromosome[1]. By contrast, autosomal imprinted genes such as *Igf2* and *H19* are monoallelically expressed according to the parent of origin[2]. The third class, randomly monoallelically transcribed autosomal genes, includes the large family of odorant-receptor genes[3] as well as genes encoding the immunoglobulins[4], T-cell receptors[5], interleukins[6,7], natural killer-cell receptors[8] and pheromone receptors[9].

All monoallelically expressed genes share the property of asynchronous replication[3,7,10], defined as one allele replicating earlier in S phase than the other allele. For most other genes, both alleles replicate synchronously at a defined portion of S phase. Asynchronous replication is an epigenetic mark that appears before transcription and may underlie the differential behavior of two alleles of identical sequence[10]. For those genes whose transcription is randomly monoallelic, the asynchronous replication is also random. The asynchronous replication seems to be established early in development before tissue-specific transcription is established[10,11] and is therefore found even in tissues in which the genes are not expressed[3]. For example, the asynchronous replication of odorant-receptor genes has been observed in all cell types analyzed, including fibroblasts and lymphocytes. The presence of asynchronous replication in a variety of cell types allowed us to compare the replication timing of diverse monoallelically expressed genes that are expressed in different cells. Given that these genes are widely dispersed across autosomes[12,13], we sought to establish the extent to which their replication asynchrony is coordinated. We focused on four autosome pairs, each containing distinct loci of randomly monoallelically expressed genes (Fig. 1*a*).

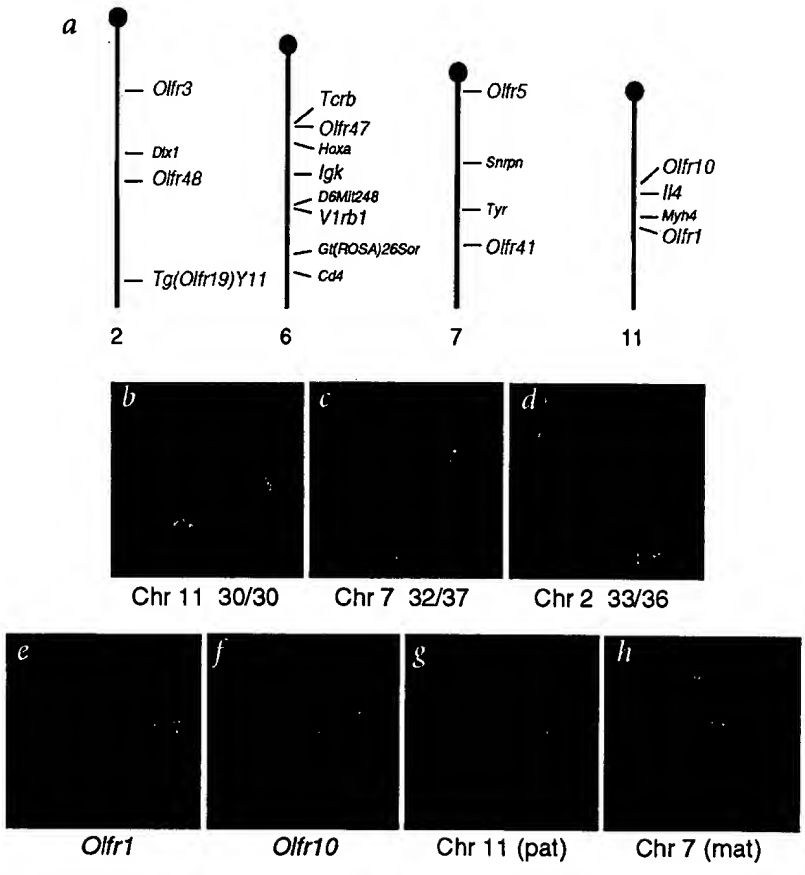


Fig. 1 Coordination of odorant-receptor asynchronous replication for individual mouse chromosome pairs. *a*, Diagram showing the relative positions of odorant-receptor genes (red) and other monoallelically expressed genes analyzed in this study (blue) along with the location of control genes (black). Centromeric ends are at the top. *b–d*, Two-color FISH analysis was done on a population of mouse embryonic fibroblasts. Blue represents DAPI staining of chromatin. *b*, Analysis of Chromosome 11. The Cy3-labeled probe (red) identifies the *Olfr1* odorant-receptor gene and the FITC-labeled probe (green) identifies *Olfr10*. The double-dot signals for the two probes in these images are on the same chromosome, indicating coordination of these two distant loci (30 of 30 cells counted). A control probe, *Myh4*, located between the *Olfr1* and *Olfr10* loci, is synchronously replicating (9% single dot–double dot pattern). *c*, Similar analysis of Chromosome 7 for two odorant-receptor genes from distinct clusters, *Olfr5* (red) and *Olfr41* (green), showing coordination (32 of 37 cells counted). Control probes between the *Olfr5* and *Olfr41* loci included the gene encoding tyrosinase, which is synchronously replicated in wild-type cells (13% single dot–double dot pattern), and the asynchronous but imprinted gene *Snrpn*. As expected, *Snrpn* did not show coordination with an odorant receptor (data not shown). *d*, Similar analysis showing that two odorant-receptor genes from distinct clusters of Chromosome 2 were coordinated (33 of 36 cells counted): *Olfr48* (red), *Olfr3* (green). A control probe, *Dlx1*, is synchronously replicating (10% single dot–double dot pattern). *e,f*, FISH analyses of line A.5 (see Supplementary Table 1 online) detected with a β-geo probe (green) that identifies the maternal Chromosome 11. Examples of cells probed with *Olfr1* (red, *e*) and *Olfr10* (red, *f*) are shown. Both of these odorant-receptor genes are maternal early-replicating in line A.5 (for data on all similar cell lines, see Supplementary Table 2 online). *g,h*, Lack of coordination between Chromosomes 7 and 11. In each case, the maternal chromosome is marked by the green probe. The maternal Chromosome 11 has a β-geo insertion and the paternal Chromosome 7 has a deletion at the tyrosinase locus. Line A.1 (see Supplementary Table 3 online) shows lack of coordination between paternally early-replicating *Olfr1* on Chromosome 11 (*g*) and maternally early-replicating *Olfr41* on Chromosome 7 (*h*).

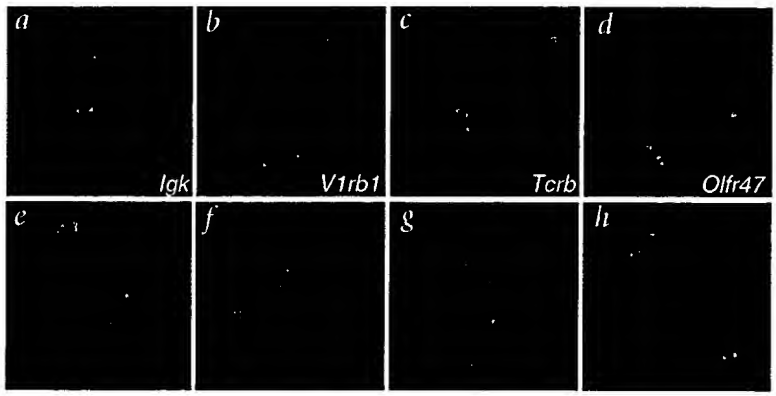


Fig. 2 A variety of monoallelically expressed genes show coordination. *a–d*, Analyses of a clonal cell line that has a marked maternal copy of Chromosome 6 (β-geo transgene, green). Probes for the *Igk* cluster (a), a large *V1R* pheromone receptor cluster (VNO-61; *b*), the *Tcrb* (c) and an odorant-receptor cluster containing *Olfr47* (*d*) are shown (red). Line F.1 has all four genes maternally early replicating. Line C.1 (see Supplementary Fig. 2 online) has all four paternally early replicating. Two control probes, *Hoxa* and *Cd4*, each showed synchronous replication (12% and 18% single dot–double dot pattern, respectively). *e*, Analyses of uncloned fibroblasts (similar to the analyses shown in Fig. 1*b–d*) also showed coordination of an odorant-receptor gene cluster on Chromosome 6 containing *Olfr47* (red) and the VNO *V1R* cluster (green; 30 of 32 cells). *f*, Same analysis as in e, but comparing coordination of *Olfr47* with *Tcrb* (33 of 34 cells). *g*, A similar population analysis showed coordination of *Il4* (red) and *Olfr10* (green) on Chromosome 11 (25 of 26 cells). *h*, Example of coordination of endogenous *Olfr48* (green) and transgenic *Tg(Olfr19)Y11* (red) on Chromosome 2 (30 of 31 cells; $P < 0.0000001$).

Asynchronous replication can be assayed by fluorescence *in situ* hybridization (FISH) analysis of interphase nuclei[14]. Replicated loci are visualized as a double-dot hybridization signal, whereas unreplicated loci are visible as a single dot. Asynchronously replicating genes present a single dot–double dot pattern in 30–40% of S-phase cells, whereas synchronously replicating genes present this pattern in roughly 10–15% of S-phase cells[14]. Although the FISH assay is only an indirect assessment of replication timing, asynchronous replication observed with this assay has been corroborated by direct measurements of replication timing (refs. 10,11; see Supplementary Note 1 and Supplementary Fig. 1 online). To assess coordination of distant loci on a given chromosome, we used two-color FISH analysis to examine two genes simultaneously and scored cells that presented a single dot–double dot signal for both genes. If the two genes are coordinated, and are replicated during a overlapping portion of S phase, the double dots for both genes should be on the same chromosome (maternal or paternal) and therefore close to each other in the nucleus. If the two genes are not coordinated, the double dots for both genes should be on the same chromosome only 50% of the time.

Using this approach we assessed the potential for coordination of asynchronous replication in wild-type primary mouse embryonic fibroblasts, analyzing two distinct odorant-receptor loci on Chromosome 11 that are 14 cM apart. Notably, we observed coordination in all of the 30 cells in which both probes presented the single dot–double dot pattern (Fig. 1*b*). Similarly, we observed coordination for two distant loci on Chromosome 7 (in 32 of 37 cells) and for two distant loci on Chromosome 2 (in 33 of 36 cells; Fig. 1*c,d*). As expected, in each case, genes between the distinct odorant-receptor loci replicated synchronously. These data indicate that odorant-receptor genes have long-range coordination of their replication asynchrony for the three autosomes examined.

We examined whether coordinated asynchronous replication of odorant-receptor genes, once established, is heritable in the progeny of a given cell. We derived clonal cell lines from embryonic and adult mice with distinguishable maternal and paternal chromosomes for Chromosomes 7 and 11. We analyzed seven cell lines for Chromosome 7 and eight cell lines for Chromosome 11. In some cell lines we consistently observed early replication of the maternal allele (Fig. 1*e*), and in the other cell lines we consistently observed early replication of the paternal allele (see Supplementary Tables 1 and 2 online). These analyses indicate that for each odorant-receptor gene, the random choice of one of the two alleles to replicate early, once established, is heritable. Analyses of these clonal cell lines also confirmed coordination of asynchronous replication along a given chromosome. For all the cell lines analyzed, both loci replicated the same parental allele early (Fig. 1*e,f* and see Supplementary Tables 1 and 2 online).

To test for genome-wide coordination, we analyzed clonal cell lines derived from mice carrying marks allowing us to distinguish the parental origins of two chromosome pairs at a time. We used FISH analysis to compare the replication timing of odorant-receptor genes on Chromosomes 6 and 11 with odorant-receptor genes on Chromosome 7 in pairwise comparisons. Odorant-receptor genes on different chromosomes were not coordinated in their replicative asynchrony; we observed all possible outcomes (Fig. 1*g,h* and see Supplementary Table 3 online). Thus, rather than genome-wide coordination, asynchronous replication of odorant-receptor genes seems to be coordinated only at the level of each chromosome pair.

To explore whether other randomly monoallelically transcribed genes are also coordinated in their replicative asynchrony, we examined clonal cell lines in which we can distinguish the maternal and paternal copies of Chromosome 6. We observed asynchronous replication of the *Igk* constant region, a V1R pheromone-receptor gene and *Tcrb* (Fig. 2*a–c*). Notably, all three of these loci were coordinated with each other as well as with an odorant-receptor gene cluster on Chromosome 6 (Fig. 2*a–d*). In some clonal cell lines the maternal alleles of all four genes replicated early (Fig. 2*a–d*), and in others the paternal alleles of all four genes replicated early (see Supplementary Fig. 2 online). Analyses of uncloned populations of cells (similar to the analyses presented in Fig. 1*b–d*) also showed coordination of genes on Chromosome 6 (Fig. 2*e,f*). Similar population analyses showed that on Chromosome 11, the gene encoding interleukin-4 (*Il4*) was coordinated with the odorant-receptor genes (Fig. 2*g*). These data, taken together, indicate that all randomly asynchronously replicated genes examined are coordinated along each chromosome.

One question arising from these observations is whether the coordination mechanisms used by different chromosomes can communicate with each other if sequences from different chromosomes are artificially placed in *cis*. We analyzed a small odorant-receptor translocation that we created artificially: a 300-kb odorant receptor–containing YAC transgene derived from Chromosome 16 that is integrated on

Chromosome 2 (*Tg(Olfr19)Y11*). We previously showed that this transgenic odorant-receptor locus undergoes asynchronous replication[15]. Two-color FISH analysis showed that the transgene was coordinated in its asynchronous replication with the endogenous odorant-receptor loci on Chromosome 2 (Fig. 2*h*), suggesting similarities in the mechanisms governing allele-specific replication timing on different chromosomes.

Here, we present data indicating that randomly monoallelically expressed genes coordinate their asynchronous replication within each chromosome pair. Scattered genes along a given chromosome are coordinated in their asynchronous replication timing in *cis*, leaving unaffected the bulk of the genes (which are synchronously replicated or, in rare instances, asynchronous but imprinted). Asynchronous replication is established early in development[10,11] and maintained in the progeny of individual cells in a clonal manner (ref. 10; Fig. 1*e–h* and see Supplementary Tables 1 and 2 online). Randomly monoallelically expressed genes are expressed in different cells of a given cell type or in different cell types. Therefore, coordination of replication timing does not imply coordination of transcription of distinct gene families and is instead a consequence of the early developmental mechanisms that establish asynchronous replication. Each gene family probably makes use of asynchronous replication (and the differences in chromatin structure that it reflects) in the complex gene regulation that characterizes these gene families. In the case of the immunoglobulin genes, we have recently shown that early replication correlates with the allele that will first undergo rearrangement and therefore provides a basis for the establishment of allelic exclusion[10].

X inactivation has been known for decades. Our data indicate that chromosome-pair non-equivalence is also found on autosomes and thus is a general, fundamental property of chromosomes that affects a large number of loci dispersed throughout the genome. The autosomal non-equivalence we observe is similar to that observed with X inactivation, except that a larger fraction of the genes on the X chromosome are affected. Other similarities in the underlying mechanisms of X inactivation and autosomal non-equivalence may emerge with further investigation.

Note: Supplementary information is available on the Nature Genetics website.

Nandita Singh[1]*, Farah A.W. Ebrahimi[1,2]*, Alexander A. Gimelbrant[1], Alexander W. Ensminger[1,2], Michael R. Tackett[1,2], Peimin Qi[1], Joost Gribnau[1] & Andrew Chess[1,2]

**These authors contributed equally to this work.*

[1]Whitehead Institute for Biomedical Research, Nine Cambridge Center, Cambridge, Massachusetts 02142, USA. [2]Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. Correspondence should be addressed to A.C. (e-mail: chess@wi.mit.edu).

Acknowledgments

We thank H. Cedar, G. Fink, D. Housman, R. Jaenisch, D. Page, C. Cowles and members of A.C.'s laboratory for discussions and comments on the manuscript; H. Higgins for manuscript preparation; G. Paradis for help with fluorescence-activated cell sorting analyses; J. Young for Chromosome 6 BACs, D. Littman for the Cd4 cosmid; R. Lane for the Olfr41 BAC; B. Holdener for the tyrosinase deletion mice; A. Dunn for the CSF2 knock-in mice; P. Soriano for the β-geo transgenic mice; N. Rosenberg for the Abelson leukemia virus–producing lines; D. MacAlpine for advice on immunoprecipitation; S. Gabriel, B. Blumenstiel, M. DeFelice and E. Winchester for help with MALDI-TOF genotyping; and S. Rozen for help with sequencing. This work was supported by grants from the U.S. National Institutes of Health (National Institute on Deafness and Other Communication Disorders) to A.C., F.A.W.E. and M.T. A. W.E. is a Howard Hughes Medical Institute Predoctoral Fellow.

Competing interests statement

The authors declare that they have no competing financial interests.

Received 6 December 2002; accepted 19 January 2003.

1. Lyon, M.F. *Nature* 320, 313 (1986).
2. Efstratiadis, A. *Curr. Opin. Genet. Dev.* 4, 265–280 (1994).
3. Chess, A., Simon, I., Cedar, H. & Axel, R. *Cell* 78, 823–834 (1994).
4. Pernis, B., Chiappino, G., Kelus, A.S. & Gell, P.G. *J. Exp. Med.* 122, 853–876 (1965).
5. Rajewsky, K. *Nature* 381, 751–758 (1996).
6. Rhoades, K.L. *et al. Curr. Biol.* 10, 789–792 (2000).
7. Hollander, G.A. *et al. Science* 279, 2118–2121 (1998).
8. Held, W., Roland, J. & Raulet, D.H. *Nature* 376, 355–358 (1995).
9. Rodriguez, I., Feinstein, P. & Mombaerts, P. *Cell* 97, 199–208 (1999).
10. Mostoslavsky, R. *et al. Nature* 414, 221–225 (2001).
11. Simon, I. *et al. Nature* 401, 929–932 (1999).
12. Young, J.M. *et al. Hum. Mol. Genet.* 11, 535–546 (2002).
13. Zhang, X. & Firestein, S. *Nat. Neurosci.* 5, 124–133 (2002).
14. Selig, S., Okumura, K., Ward, D.C. & Cedar, H. *EMBO J.* 11, 1217–1225 (1992).
15. Ebrahimi, F.A., Edmondson, J., Rothstein, R. & Chess, A. *Dev. Dyn.* 217, 225–231 (2000).

Commentary

Cancer epigenetics takes center stage

Andrew P. Feinberg

Institute of Genetic Medicine, Departments of Medicine, Molecular Biology and Genetics, and Oncology, Johns Hopkins University School of Medicine, 1064 Ross, 720 Rutland Avenue, Baltimore, MD 21205

Next year will mark 20 years since I developed a Southern blot showing altered DNA methylation in cancer. This discovery (1) was met with some skepticism, primarily because it was thought that aberrant methylation in cancer was an epiphenomenon, somehow linked to a generalized disruption of gene regulation in cancer cells and arising after the cancer, rather than playing a causal role itself. This essay will address how cancer epigenetics has overcome these objections, and a report in this issue by Nakagawa *et al.* (2) adds significantly to this argument.

Epigenetics is defined as modifications of the genome, heritable during cell division, that do not involve a change in the DNA sequence. Examples include methylation induced premeiotically in *Ascobolus*, repeat-induced gene silencing and paramutation in plants, mating type silencing and telomere silencing in yeast, position effect variegation in *Drosophila*, and genomic imprinting in mammals and flowering plants. There are several features that distinguish epigenetics from conventional genetic mechanisms: reversibility; position effects, i.e., the ability to act over unexpected distances larger than a single gene; apparent mutations at unexpectedly high frequency; and the involvement of gene domains. One common thread to most epigenetic phenomena is DNA methylation, a covalent modification of the C5 position of cytosine. This methylation pattern is stably maintained at CpG dinucleotides by a family of DNA methyltransferases that recognize hemimethylated CpG dinucleotides after DNA replication. DNA methyltransferases belong to multiprotein complexes, and they contain sequence motifs for multiple such interactions, including interactions with chromatin components, some of which have been directly identified (3). Another common thread to epigenetics is a link to transcriptional regulation, generally involving gene silencing. DNA methylation in particular is generally but not exclusively linked to transcriptional silencing, including methylation induced premeiotically, paramutation, and mammalian gene silencing.

Epigenetic alterations in cancer include global hypomethylation (4), hypomethylation of individual genes (1), and hypermethylation of CpG islands (5), CpG-rich sequences in the promoters of housekeeping genes that are generally protected from methylation. This hypermethylation may lead to aberrant silencing of tumor suppressor genes (6). In addition, we and others have discovered loss of imprinting (LOI) in cancer (7, 8). Genomic imprinting, the subject of the report by Nakagawa *et al.* (2), is an epigenetic modification of a specific parental allele of a gene, or the chromosome on which it resides, in the gamete or zygote, leading to differential expression of the two alleles of the gene in somatic cells of the offspring. LOI involves loss of the normal pattern of expression of a specific parental allele, and in cancer it can lead to activation of growth-promoting imprinted genes such as insulin-like growth factor II (*IGF2*) (7, 8), as well as silencing of potential tumor suppressor genes such as $p57^{KIP2}$ (9) and *ARH1* (10).

Furthermore, we found that LOI can occur in the *normal* colonic mucosa of colorectal cancer patients with LOI in their tumors (11), overcoming the objection that epigenetic alterations are simply late consequences of neoplasia. This LOI was linked to cases showing microsatellite instability (MSI) in the tumors (11). MSI is a form of genetic instability found in patients with hereditary nonpolyposis colorectal cancer and caused by defects in DNA mismatch repair (12, 13). MSI occurs much more commonly in sporadic nonfamilial colon cancer, affecting about 25% of such patients. However, these patients do not have mutations in mismatch repair genes (14). One potential cause of MSI in these sporadic cancers is hypermethylation and epigenetic silencing of the hMLH1 mismatch repair gene (15), a target of conventional mutations in hereditary nonpolyposis colorectal cancer (16, 17). LOI of *IGF2* has also been previously linked to increased methylation in embryonal and other tumors, specifically at a CpG island that represents a differentially methylated region (DMR) upstream of the maternal H19 gene; methylation of the DMR in turn regulates the silencing of the IGF2 gene on the same chromosome (18, 19).

Nakagawa *et al.* (2) now confirm the original study of Cui *et al.* that LOI occurs in both tumor and normal tissue of patients with MSI-positive colorectal cancer (11), a result also confirmed earlier by Nishihara *et al.* (20). Furthermore, Nakagawa *et al.* find that LOI was not a feature of tumors in patients with a germline mismatch repair gene mutation (2), indicating that LOI is not merely a consequence of a mismatch repair defect. The study of Nakagawa *et al.* (2) also helps to close the circle among cancer, LOI, and DNA methylation, by demonstrating directly that sequences within the *H19* DMR are specifically methylated in these tumors. Not surprisingly, these same patients show hypermethylation of other CpG islands throughout the genome, as generalized hypermethylation and MSI have been previously linked (21, 41). Methylation of the H19 DMR was also observed in the matched normal tissue of patients with LOI (2), although it should be noted that this occurred at a lower frequency than LOI in these tissues, and only partial methylation was observed.

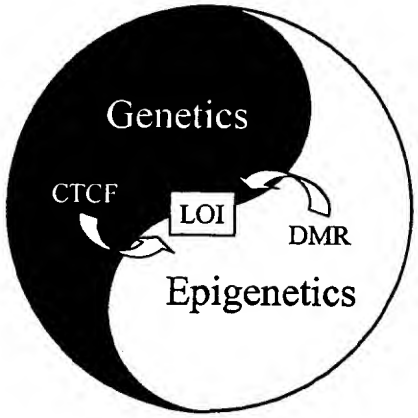


Fig. 1. The interrelationship of cancer genetics and epigenetics. Rather than a traditional Venn diagram, cancer genetics and epigenetics are drawn as a yin-yang. For example, loss of imprinting might be caused by genetic disruption of CTCF or by altered methylation of the DMR (illustrated). Other examples of overlap or in which the distinction between genetics and epigenetics are blurred are provided in the text.

See companion article on page 591.

Furthermore, generalized hypermethylation of CpG islands is *not* present in the normal tissue of these patients (2), consistent with the idea that LOI precedes a generalized disruption of CpG island methylation (11).

The present study (2) also offers an intriguing mechanistic hypothesis to explain the relationship between *H19* DMR methylation and LOI in these patients, as the methylated nucleotides include those to which the chromatin insulator CTCF has been shown to bind specifically in regulating genomic imprinting. Whether the partial methylation seen in their study is sufficient to disrupt CTCF binding remains to be proven. Nevertheless, the study calls attention to this remarkable highly conserved multifunctional protein, first discovered by Lobanenkov and colleagues as a multivalent transcription factor (22–24) that also serves as a chromatin insulator. In this capacity, CTCF binds specifically to the *H19* DMR *in vivo* (25) and *in vitro* (26–28) when it is unmethylated, separating *IGF2* from its enhancer and allowing monoallelic DMR methylation-dependent expression of *IGF2* (25–28). Lobanenkov and colleagues (including us) have found mutations in *CTCF* in diverse tumors, that selectively impair binding to target sequences, altering the functional spectrum of the protein and, as well as methylation of CTCF binding sites in tumors (G. N. Filippova, D. I. Loukinov, E. M. Pugacheva, J. E. Ulmer, J. M. Moore, Y. J. Hu, H. Moon, J. Breen, C.-F. Qi, P. E. Grundy, *et al.*, unpublished work), suggesting a general role for CTCF in cancer.

The potential link to CTCF suggested by this study also calls our attention to the link among DNA methylation, epigenetics, and chromatin. Boveri, the father of cancer genetics, described a generalized disturbance of chromatin that distinguished cancer cells from normal cells (30). Although his writing preceded our understanding of DNA and thus he was not truly distinguishing genetics and epigenetics, his thoughts were based on his observation of widespread disruption of chromosomal organization and nuclear structure, akin to what we might call chromatin today. In this regard, a gem in the recent studies of CTCF was the observation of Ohlsson and colleagues that CTCF binding may depend on nucleosome phasing (31), suggesting that imprinting and methylation may both ultimately be bound to the assembly of DNA into organized structures.

A clue to the link between MSI and epigenetics may be provided by another sometimes overlooked common thread in epigenetics, namely DNA replication. For example, repeat-induced gene silencing is thought to be propagated through hemimethylated intermediates during DNA replication; silencing in yeast depends on an origin of replication complex; and mating type silencing, position effect variegation, and genomic imprinting are all linked to delayed replication timing. Given that MSI is attributable to defective replication-linked mismatch repair, necessitating that the cell distinguishes between parent and daughter strands, all of these phenomena may involve disrupted chromatin.

The studies of Cui *et al.* (11), Nishihara *et al.* (20), and Nakagawa *et al.* (2) suggest a new and provocative view of the timing of epigenetic changes in cancer. It should be noted that LOI was observed in these studies generally throughout the colon, not just at the site of the tumor. Although colorectal cancer in particular involves a series of genetic alterations in the evolution of an advanced metastatic tumor, elegantly described by Vogelstein (32) and others, perhaps the likelihood of developing clinical cancer when a mutation arises depends in part on a preexisting epigenetic defect affecting much or all of the normal colonic mucosa. Why would this be? Studies of transgenic mice with constitutive biallelic expression of *IGF2*, comparable to LOI, show reduced apoptosis and increased tumor formation on introduction of an oncogenic transgene (33, 34). Perhaps preexisting LOI alters the balance between growth and apoptosis when conventional oncogenic mutations arise in the colon. Although this idea is admittedly speculative, these studies (2, 11, 20) nevertheless demonstrate such a field defect in nontumor tissue, and from an epidemiological perspective, what distinguishes a cancer patient from a noncancer patient may be such epigenetic alterations involving DNA methylation and imprinting. This hypothesis will require direct confirmation that the epigenetic alterations temporally precede the genetic changes and the tumors themselves, a subject of intense current clinical study. Such studies may have profound clinical significance, because they might offer the opportunity for detection of large numbers of patients in the general population at increased risk of colorectal cancer, and the eventual possibility of enhanced surveillance or even chemoprevention in such patients.

I conclude by noting that the distinction between cancer genetics and epigenetics has blurred considerably in recent years (Fig. 1). Many conventional "genetic" mechanisms directly affect proteins that regulate chromatin, such as the rearrangement of the trithorax family member ALL in common childhood leukemia (29, 35), mutation of the chromatin remodeling complex core member Snf5 in rhabdoid tumors (36), and rearrangement of a candidate histone acetyltransferase in acute myeloid leukemia (37). In addition, physical interactions have been found between critical tumor genes and chromatin proteins, such as Rb and histone deactylase (38, 39), and the promyelocytic leukemia nuclear body protein P100 with heterochromatin protein HP-1 (40), whose *Drosophila* homologue is also a suppressor of position effect variegation. While less is known of the mechanism of epigenetic alterations in cancer, CTCF itself offers a striking example of convergence of traditional genetics and epigenetics, as the same protein regulates genomic imprinting as well as the "traditional" oncogene c-*myc* (23). We may well find that the geneticists and the epigeneticists converge on Boveri's definition of cancer as a disease of "the chromatin."

I thank R. Ohlsson, V. Lobanenkov, H. Cui, J. Ravenel, and E. Niemitz for thoughtful comments. This work was supported by National Institutes of Health Grant CA65145.

1. Feinberg, A. P. & Vogelstein, B. (1983) *Nature (London)* **301**, 89–92.
2. Nakagawa, H., Chadwick, R. B., Peltomäki, P., Plass, C., Nakamura, Y. & de la Chapelle, A. (2001) *Proc. Natl. Acad. Sci. USA* **98**, 591–596. (First Published December 19, 2000; 10.1073/pnas.011528698)
3. Robertson, K. D., Ait-Si-Ali, S., Yokochi, T., Wade, P. A., Jones, P. L. & Wolffe, A. P. (2000) *Nat. Genet.* **25**, 338–342.
4. Feinberg, A. P., Gehrke, C. W., Kuo, K. C. & Ehrlich, M. (1988) *Cancer Res.* **48**, 1159–1161.
5. Baylin, S. B., Hoppener, J. W., de Bustros, A., Steenbergh, P. H., Lips, C. J. & Nelkin, B. D. (1986) *Cancer Res.* **46**, 2917–2922.
6. Jones, P. A. & Laird, P. W. (1999) *Nat. Genet.* **21**, 163–167.
7. Rainier, S., Johnson, L. A., Dobry, C. J., Ping, A. J., Grundy, P. E. & Feinberg, A. P. (1993) *Nature (London)* **362**, 747–749.
8. Ogawa, O., Eccles, M. R., Szeto, J., McNoe, L. A., Yun, K., Maw, M. A., Smith, P. J. & Reeve, A. E. (1993) *Nature (London)* **362**, 749–751.
9. Thompson, J. S., Reese, K. J., DeBaun, M. R., Perlman, E. J. & Feinberg, A. P. (1996) *Cancer Res.* **56**, 5723–5727.
10. Yu, Y., Xu, F., Peng, H., Fang, X., Zhao, S., Li, Y., Cuevas, B., Kuo, W. L., Gray, J. W., Siciliano, M., *et al.* (1999) *Proc. Natl. Acad. Sci. USA* **96**, 214–219.
11. Cui, H., Horon, I. L., Ohlsson, R., Hamilton, S. R. & Feinberg, A. P. (1998) *Nat. Med.* **4**, 1276–1280.
12. Thibodeau, S. N., Bren, G. & Schaid, D. (1993) *Science* **260**, 816–819.
13. Parsons, R., Li, G. M., Longley, M. J., Fang, W. H., Papadopoulos, N., Jen, J., de la Chapelle, A., Kinzler, K. W., Vogelstein, B. & Modrich, P. (1995) *Science* **268**, 738–740.

COMMENTARY

14. Boland, C. R., Thibodeau, S. N., Hamilton, S. R., Sidransky, D., Eshleman, J. R., Burt, R. W., Meltzer, S. J., Rodriguez-Bigas, M. A., Fodde, R., Ranzani, G. N., *et al.* (1998) *Cancer Res.* **58**, 5248–5257.
15. Herman, J. G., Umar, A., Polyak, K., Graff, J. R., Ahuja, N., Issa, J. P., Markowitz, S., Willson, J. K., Hamilton, S. R., Kinzler, K. W., *et al.* (1998) *Proc. Natl. Acad. Sci. USA* **95**, 6870–6875.
16. Papadopoulos, N., Nicolaides, N. C., Wei, Y. F., Ruben, S. M., Carter, K. C., Rosen, C. A., Haseltine, W. A., Fleischmann, R. D., Fraser, C. M., Adams, M. D., *et al.* (1994) *Science* **263**, 1625–1629.
17. Bronner, C. E., Baker, S. M., Morrison, P. T., Warren, G., Smith, L. G., Lescoe, M. K., Kane, M., Earabino, C., Lipford, J., Lindblom, A., *et al.* (1994) *Nature (London)* **368**, 258–261.
18. Steenman, M. J. C., Rainier, S., Dobry, C. J., Grundy, P., Horon, I. L. & Feinberg, A. P. (1994) *Nat. Genet.* **7**, 433–439.
19. Moulton, T., Crenshaw, T., Hao, Y., Moosikasuwan, J., Lin, N., Dembitzer, F., Hensle, T., Weiss, L., McMorrow, L., Loew, T., *et al.* (1994) *Nat. Genet.* **7**, 440–447.
20. Nishihara, S., Hayashida, T., Mitsuya, K., Schulz, T. C., Ikeguchi, M., Kaibara, N. & Oshimura, M. (2000) *Int. J. Oncol.* **17**, 317–322.
21. Toyota, M. & Issa, J. P. (1999) *Semin. Cancer Biol.* **9**, 349–357.
22. Lobanenkov, V. V., Nicolas, R. H., Plumb, M. A., Wright, C. A. & Goodwin, G. H. (1986) *Eur. J. Biochem.* **159**, 181–188.
23. Klenova, E. M., Nicolas, R. H., Paterson, H. F., Carne, A. F., Heath, C. M., Goodwin, G. H., Neiman, P. E. & Lobanenkov, V. V. (1993) *Mol. Cell. Biol.* **13**, 7612–7624.
24. Filippova, G. N., Fagerlie, S., Klenova, E. M., Myers, C., Dehner, Y., Goodwin, G., Neiman, P. E., Collins, S. J. & Lobanenkov, V. V. (1996) *Mol. Cell. Biol.* **16**, 2802–2813.
25. Kanduri, C., Pant, V., Loukinov, D., Pugacheva, E., Qi, C. F., Wolffe, A., Ohlsson, R. & Lobanenkov, V. V. (2000) *Curr. Biol.* **10**, 853–856.
26. Hark, A. T., Schoenherr, C. J., Katz, D. J., Ingram, R. S., Levorse, J. M. & Tilghman, S. M. (2000) *Nature (London)* **405**, 486–489.
27. Bell, A. C. & Felsenfeld, G. (2000) *Nature (London)* **405**, 482–485.
28. Szabo, P., Tang, S. H., Rentsendorj, A., Pfeifer, G. P. & Mann, J. R. (2000) *Curr. Biol.* **10**, 607–610.
29. Tkachuk, D. C., Kohler, S. & Cleary, M. L. (1992) *Cell* **71**, 691–700.
30. Boveri, T. (1929) *The Origin of Malignant Tumors* (Williams & Wilkins, Baltimore).
31. Kanduri, C., Holmgren, C., Pilartz, M., Franklin, G., Kanduri, M., Liu, L., Ginjala, V., Ulleras, E., Mattsson, R. & Ohlsson, R. (2000) *Curr. Biol.* **10**, 449–457.
32. Fearon, E. R. & Vogelstein, B. (1990) *Cell* **61**, 759–767.
33. Christofori, G., Naik, P. & Hanahan, D. (1994) *Nature (London)* **369**, 414–418.
34. Christofori, G., Naik, P. & Hanahan, D. (1995) *Nat. Genet.* **10**, 196–201.
35. Gu, Y., Nakamura, T., Alder, H., Prasad, R., Canaani, O., Cimino, G., Croce, C. M. & Canaani, E. (1992) *Cell* **71**, 701–708.
36. Versteege, I., Sevenet, N., Lange, J., Rousseau-Merck, M. F., Ambros, P., Handgretinger, R., Aurias, A. & Delattre, O. (1998) *Nature (London)* **394**, 203–206.
37. Borrow, J., Stanton, V. P., Jr., Andresen, J. M., Becher, R., Behm, F. G., Chaganti, R. S., Civin, C. I., Disteche, C., Dube, I., Frischauf, A. M., *et al.* (1996) *Nat. Genet.* **14**, 33–41.
38. Magnaghi-Jaulin, L., Groisman, R., Naguibneva, I., Robin, P., Lorain, S., Le Villain, J. P., Troalen, F., Trouche, D. & Harel-Bellan, A. (1998) *Nature (London)* **391**, 601–605.
39. Brehm, A., Miska, E. A., McCance, D. J., Reid, J. L., Bannister, A. J. & Kouzarides, T. (1998) *Nature (London)* **391**, 597–601.
40. Seeler, J. S., Marchio, A., Sitterlin, D., Transy, C. & Dejean, A. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 7316–7321.
41. Lengauer, C., Kinzler, K. W. & Vogelstein, B. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 2545–2550.